Framework For Comparing Error Covariance Estimation Schemes

Michael Sitwell, Environment and Climate Change Canada

Mathematical Approaches of Atmospheric Constituents Data Assimilation and Inverse Modeling workshop

Banff, March 23, 2023

(Richard Ménard, ECCC)

Environnement Canada Environment Canada

Canada

Introduction

- Specification of error covariances can impact the quality of the analyses
  - Background error covariances $\mathbf{B}$
  - Observation error covariances $\mathbf{R}$

- Difficult to accurately specifying error covariances $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{R}}$
  - Helpful to have methods to diagnose and/or tune the error covariances

- For simplicity, we focus on tuning variances only

$$\widetilde{\mathbf{B}}' = s_\mathrm{b}\widetilde{\mathbf{B}} \qquad \widetilde{\mathbf{R}}' = s_\mathrm{o}\widetilde{\mathbf{R}}$$

Observation-Space Residuals

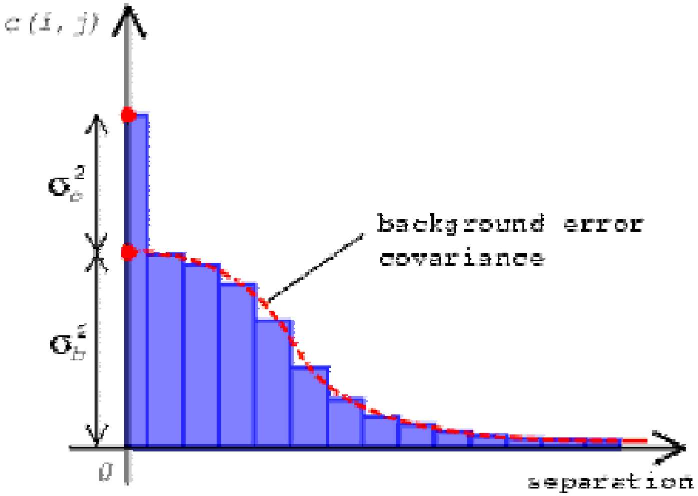
- The innovation vector $\mathbf{d}$ does not depend on the true variables

$$\mathbf{d} = \underset{\substack{\uparrow \\ \text{observations}}}{\mathbf{y}} - H(\underset{\substack{\uparrow \\ \text{background}}}{\mathbf{x}_\mathrm{b}}) \approx \boldsymbol{\epsilon}_\mathrm{o} - \underset{\substack{\uparrow \\ \text{linear obs.} \\ \text{operator}}}{\mathbf{H}}\boldsymbol{\epsilon}_\mathrm{b} \qquad \mathbf{D} = \mathrm{cov}(\mathbf{d}, \mathbf{d}) = \mathbf{HBH}^\mathrm{T} + \mathbf{R}$$

- However, $\mathbf{B}$ and $\mathbf{R}$ are not separate in $\mathbf{D}$
  - Want modelled quantities $\widetilde{\mathbf{B}}$ and $\widetilde{\mathbf{R}}$

Hollingsworth and Lönnberg

- Assumes we can attribute all spatial correlations in the innovation to the background
  - $\tilde{\mathbf{R}}$ assumes diagonal
  - $\tilde{\mathbf{B}}$ terms are fit to the observed innovation at nonzero spatial separation
  - Extrapolation to zero spatial separation



Bouttier, F., and P. Courtier. "Data assimilation concepts and methods March 1999." *Meteorological training course lecture series. ECMWF* 718 (2002): 59.

The Desroziers and Ivanov 2001 Method (DI01)

innovation covariance consistency $\mathrm{E}[\mathbf{D}] = \widetilde{\mathbf{D}}$

- For variational assimilation systems

$$\mathrm{E}[J_\mathrm{b}(\mathbf{x}_\mathrm{a})] = \tfrac{1}{2}\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}^{-1}\mathrm{E}[\mathbf{D}]\widetilde{\mathbf{D}}^{-1}] \xrightarrow{\mathrm{E}[\mathbf{D}]=\widetilde{\mathbf{D}}} \tfrac{1}{2}\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}^{-1}]$$

$$\mathrm{E}[J_\mathrm{o}(\mathbf{x}_\mathrm{a})] = \tfrac{1}{2}\mathrm{Tr}[\widetilde{\mathbf{R}}\widetilde{\mathbf{D}}^{-1}\mathrm{E}[\mathbf{D}]\widetilde{\mathbf{D}}^{-1}] \xrightarrow{\mathrm{E}[\mathbf{D}]=\widetilde{\mathbf{D}}} \tfrac{1}{2}\mathrm{Tr}[\widetilde{\mathbf{R}}\widetilde{\mathbf{D}}^{-1}]$$

- Iterative scheme:

$$(s_\mathrm{b}^\mathrm{DI01})_{i+1} = \frac{\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}_i\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}_i^{-1}\mathbf{D}\widetilde{\mathbf{D}}_i^{-1}]}{\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}_i\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}_i^{-1}]} \qquad (s_\mathrm{o}^\mathrm{DI01})_{i+1} = \frac{\mathrm{Tr}[\widetilde{\mathbf{R}}_i\widetilde{\mathbf{D}}_i^{-1}\mathbf{D}\widetilde{\mathbf{D}}_i^{-1}]}{\mathrm{Tr}[\widetilde{\mathbf{R}}_i\widetilde{\mathbf{D}}_i^{-1}]}$$

- Has been used in a NWP assimilation system, but difficult to implement due to high computational cost

The Desroziers et al. 2005 Method (D05)

$$\mathrm{E}[(\mathrm{H}(\mathbf{x}_\mathrm{a}) - \mathrm{H}(\mathbf{x}_\mathrm{b}))\mathbf{d}^\mathrm{T}] = \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}^{-1}\mathrm{E}[\mathbf{D}] \xrightarrow{\mathrm{E}[\mathbf{D}]=\widetilde{\mathbf{D}}} \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T}$$

$$\mathrm{E}[(\mathbf{y} - \mathrm{H}(\mathbf{x}_\mathrm{a}))\mathbf{d}^\mathrm{T}] = \widetilde{\mathbf{R}}\widetilde{\mathbf{D}}^{-1}\mathrm{E}[\mathbf{D}] \xrightarrow{\mathrm{E}[\mathbf{D}]=\widetilde{\mathbf{D}}} \widetilde{\mathbf{R}}$$

- Iterative scheme:

$$(s_\mathrm{b}^\mathrm{D05})_{i+1} = \frac{\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}_i\mathbf{H}^\mathrm{T}\widetilde{\mathbf{D}}_i^{-1}\mathbf{D}]}{\mathrm{Tr}[\mathbf{H}\widetilde{\mathbf{B}}_i\mathbf{H}^\mathrm{T}]} \qquad (s_\mathrm{o}^\mathrm{D05})_{i+1} = \frac{\mathrm{Tr}[\widetilde{\mathbf{R}}_i\widetilde{\mathbf{D}}_i^{-1}\mathbf{D}]}{\mathrm{Tr}[\widetilde{\mathbf{R}}_i]}$$

- Typically much less computationally demanding than DI01

Method Comparisons

- Different methods can produce very different results
- For DI01 and D05, not evident exactly how **B** and **R** are being separated
- Mathematical formalism needed for:
  - Direct comparison between methods
    - Why do they work?
  - Understanding different regimes for each method:
    - When do each method give reasonable results or fail?

Filtering and the Analysis

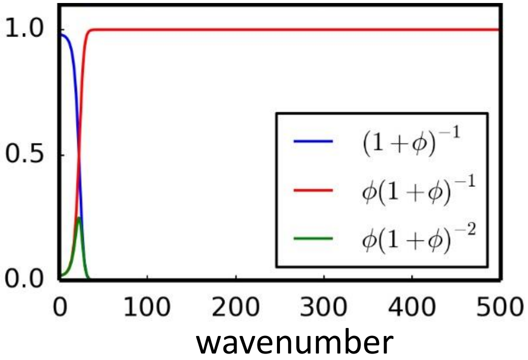
- Transform to basis that simultaneously diagonalizes and $\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$ and $\widetilde{\mathbf{R}}$

$$\phi = \frac{\text{spectra of } \widetilde{\mathbf{R}}}{\text{spectra of } \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}}$$

- Construct operators:

$\tilde{\mathbf{F}}$ that dampens modes prominent in $\widetilde{\mathbf{R}}$ as compared to $\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$

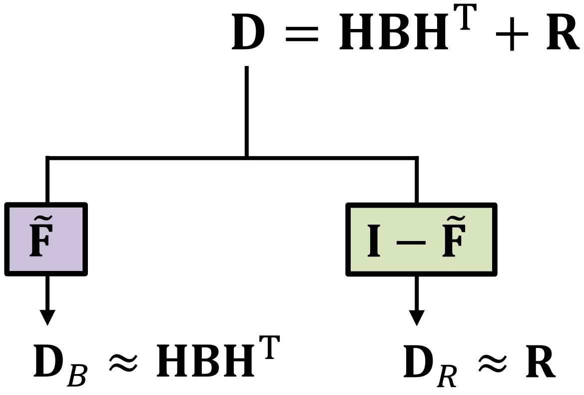
$\mathbf{I} - \tilde{\mathbf{F}}$ that dampens modes prominent in $\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}$ as compared $\widetilde{\mathbf{R}}$



1D periodic domain of length $L = 40{,}000$ km, observations $\Delta x = 40$ km apart. $L_{\mathrm{b}} = 800$ km and $L_{\mathrm{o}} = 0$

$$\mathbf{H}\mathbf{x}_{\mathrm{a}} = \left(\mathbf{I} - \tilde{\mathbf{F}}\right)\mathbf{H}\mathbf{x}_{\mathrm{b}} + \tilde{\mathbf{F}}\mathbf{y} = \mathbf{H}\mathbf{x}_{\mathrm{b}} + \tilde{\mathbf{F}}\mathbf{d}$$

Filtering of Error Covariances



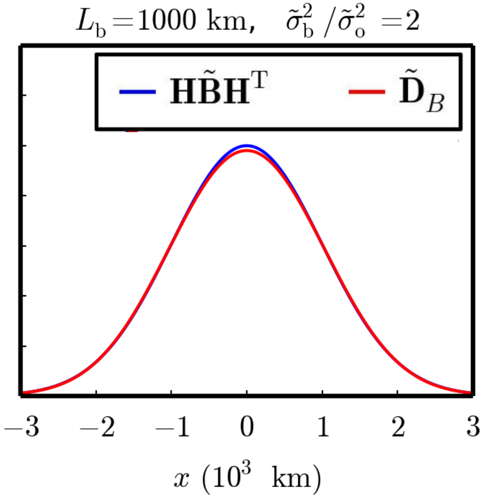
$$\mathbf{D} = \mathbf{HBH}^{\mathrm{T}} + \mathbf{R}$$

$\tilde{\mathbf{F}}$

$\mathbf{I} - \tilde{\mathbf{F}}$

$\mathbf{D}_B \approx \mathbf{HBH}^{\mathrm{T}}$

$\mathbf{D}_R \approx \mathbf{R}$



observed covariances

$$\mathbf{D}_B \equiv \tilde{\mathbf{F}}\mathbf{D}\tilde{\mathbf{F}}^{\mathrm{T}}$$

$$\mathbf{D}_R \equiv (\mathbf{I} - \tilde{\mathbf{F}})\mathbf{D}(\mathbf{I} - \tilde{\mathbf{F}})^{\mathrm{T}}$$

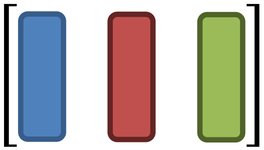
modelled covariances

$$\tilde{\mathbf{D}}_B \equiv \tilde{\mathbf{F}}\tilde{\mathbf{D}}\tilde{\mathbf{F}}^{\mathrm{T}}$$

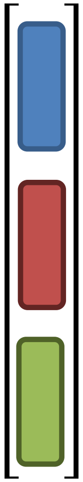
$$\tilde{\mathbf{D}}_R \equiv (\mathbf{I} - \tilde{\mathbf{F}})\tilde{\mathbf{D}}(\mathbf{I} - \tilde{\mathbf{F}})^{\mathrm{T}}$$

Vectorization of Matrices

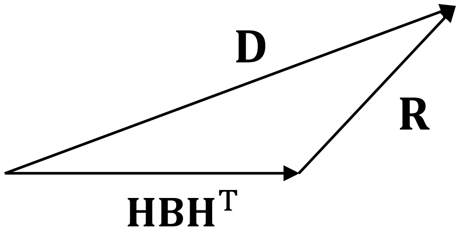
$\mathbf{A} =$



$\mathrm{vec}(\mathbf{A}) =$



$$\mathbf{D} = \mathbf{HBH}^{\mathrm{T}} + \mathbf{R}$$



Frobenius inner product: $\langle \mathbf{A}, \mathbf{B} \rangle = \mathrm{vec}(\mathbf{A}) \cdot \mathrm{vec}(\mathbf{B}) = \mathrm{Tr}[\mathbf{A}^{\mathrm{T}}\mathbf{B}]$

Linear Least-Squares Solution

$$\mathbf{y} = \mathbf{M}x + \boldsymbol{\epsilon}, \quad \mathrm{cov}(\boldsymbol{\epsilon}, \boldsymbol{\epsilon}) = \mathbf{C}$$

linear least-squares solution is $\hat{x} = \frac{\mathbf{M}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{y}}{\mathbf{M}^{\mathrm{T}}\mathbf{C}^{-1}\mathbf{M}}$

$$s_{\mathrm{o}}^{\mathrm{DI01}} = \frac{\overbrace{\mathrm{vec}\left(\widetilde{\mathbf{D}}_R\right)}^{\text{model}}{}^{\mathrm{T}}\,\overbrace{(\widetilde{\mathbf{D}}_R \otimes \widetilde{\mathbf{R}})}^{\text{covariance}}{}^{-1}\,\overbrace{\mathrm{vec}(\mathbf{D}_R)}^{\text{observed data}}}{\mathrm{vec}\left(\widetilde{\mathbf{D}}_R\right)^{\mathrm{T}}(\widetilde{\mathbf{D}}_R \otimes \widetilde{\mathbf{R}})^{-1}\mathrm{vec}(\widetilde{\mathbf{D}}_R)}$$

Linear Least-Squares Solution

$$\boxed{J_\mathrm{o} = \frac{1}{2}\|\mathbf{d}-\mathbf{H}\Delta\mathbf{x}\,\|^2_{\tilde{\mathbf{R}}^{-1}}}$$

$$J_R^{\mathrm{DI01}} \equiv \frac{1}{2}\left\|s_\mathrm{o}\tilde{\mathbf{D}}_R-\mathbf{D}_R\right\|^2_{(\tilde{\mathbf{D}}_R\otimes\tilde{\mathbf{R}})^{-1}}$$

minimize
w.r.t.
s_o

$$s_\mathrm{o}^{\mathrm{DI01}} = \frac{\overbrace{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}}^{\text{model}}\overbrace{(\tilde{\mathbf{D}}_R\otimes\tilde{\mathbf{R}})^{-1}}^{\text{covariance}}\overbrace{\mathrm{vec}(\mathbf{D}_R)}^{\text{observed data}}}{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}(\tilde{\mathbf{D}}_R\otimes\tilde{\mathbf{R}})^{-1}\mathrm{vec}(\tilde{\mathbf{D}}_R)}$$

$$J_B^{\mathrm{DI01}} \equiv \frac{1}{2}\left\|s_\mathrm{b}\tilde{\mathbf{D}}_B - \mathbf{D}_B\right\|^2_{(\tilde{\mathbf{D}}_B \otimes (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^\mathrm{T}))^{-1}}$$

minimize
w.r.t.
s_b

$$s_\mathrm{b}^{\mathrm{DI01}} = \frac{\mathrm{vec}\left(\tilde{\mathbf{D}}_B\right)^\mathrm{T}(\tilde{\mathbf{D}}_B \otimes (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^\mathrm{T}))^{-1}\mathrm{vec}(\mathbf{D}_B)}{\mathrm{vec}\left(\tilde{\mathbf{D}}_B\right)^\mathrm{T}(\tilde{\mathbf{D}}_B \otimes (\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^\mathrm{T}))^{-1}\mathrm{vec}(\tilde{\mathbf{D}}_B)}$$

$$J_R^{\mathrm{DI01}} \equiv \frac{1}{2}\left\|s_\mathrm{o}\tilde{\mathbf{D}}_R - \mathbf{D}_R\right\|^2_{(\tilde{\mathbf{D}}_R \otimes \tilde{\mathbf{R}})^{-1}}$$

minimize
w.r.t.
s_o

$$s_\mathrm{o}^{\mathrm{DI01}} = \frac{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}(\tilde{\mathbf{D}}_R \otimes \tilde{\mathbf{R}})^{-1}\mathrm{vec}(\mathbf{D}_R)}{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}(\tilde{\mathbf{D}}_R \otimes \tilde{\mathbf{R}})^{-1}\mathrm{vec}(\tilde{\mathbf{D}}_R)}$$

$$J_B^{\mathrm{D05}} \equiv \frac{1}{2}\left\|s_\mathrm{b}\tilde{\mathbf{D}}_B - \mathbf{D}_B\right\|^2_{(\tilde{\mathbf{B}}_B \otimes \mathbf{I})^{-1}}$$

minimize
w.r.t.
s_b

$$s_\mathrm{b}^{\mathrm{D05}} = \frac{\mathrm{vec}\left(\tilde{\mathbf{D}}_B\right)^\mathrm{T}(\tilde{\mathbf{B}}_B \otimes \mathbf{I})^{-1}\mathrm{vec}(\mathbf{D}_B)}{\mathrm{vec}\left(\tilde{\mathbf{D}}_B\right)^\mathrm{T}(\tilde{\mathbf{B}}_B \otimes \mathbf{I})^{-1}\mathrm{vec}(\tilde{\mathbf{D}}_B)}$$

$$J_R^{\mathrm{D05}} \equiv \frac{1}{2}\left\|s_\mathrm{o}\tilde{\mathbf{D}}_R - \mathbf{D}_R\right\|^2_{(\tilde{\mathbf{R}}_R \otimes \mathbf{I})^{-1}}$$

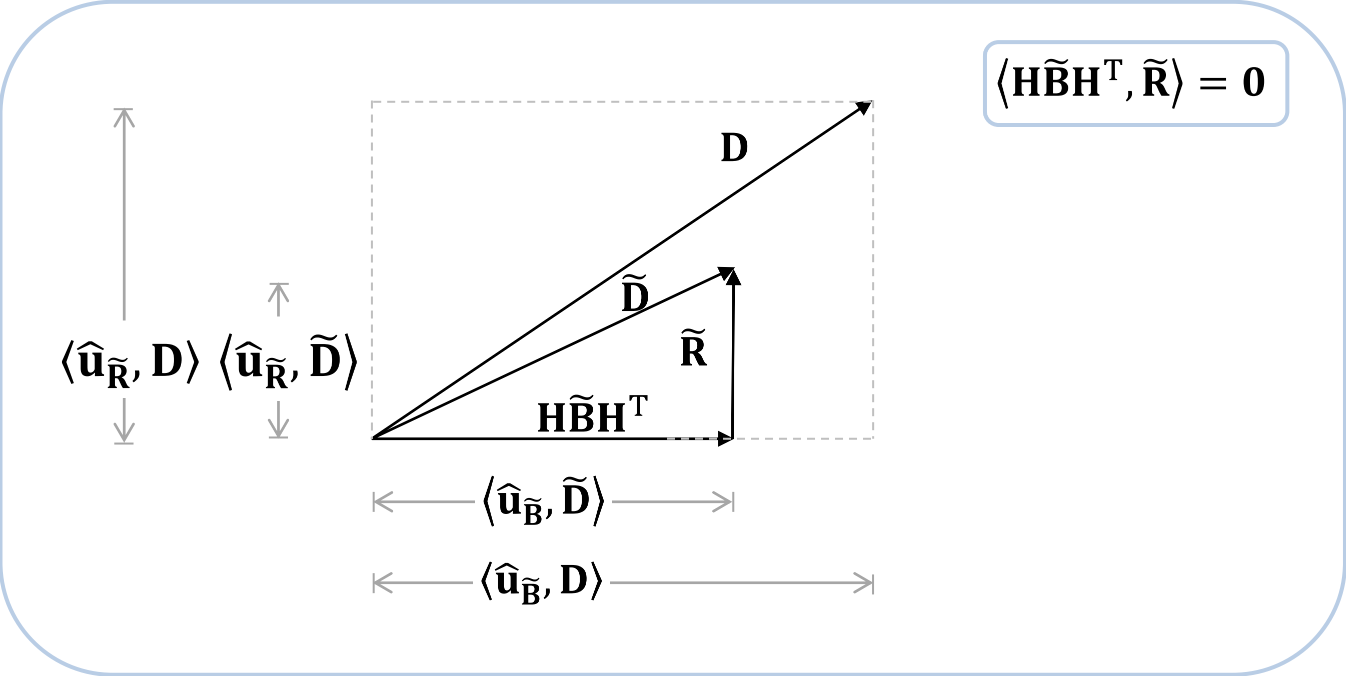
minimize
w.r.t.
s_o

$$s_\mathrm{o}^{\mathrm{D05}} = \frac{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}(\tilde{\mathbf{R}}_R \otimes \mathbf{I})^{-1}\mathrm{vec}(\mathbf{D}_R)}{\mathrm{vec}\left(\tilde{\mathbf{D}}_R\right)^\mathrm{T}(\tilde{\mathbf{R}}_R \otimes \mathbf{I})^{-1}\mathrm{vec}(\tilde{\mathbf{D}}_R)}$$

Geometric Interpretations

$$s_{\mathrm{b}}^{\mathrm{DI01}} = \frac{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{B}}}, \mathbf{D} \rangle_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}}}{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{B}}}, \tilde{\mathbf{D}} \rangle_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}}} \qquad s_{\mathrm{o}}^{\mathrm{DI01}} = \frac{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{R}}}, \mathbf{D} \rangle_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}}}{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{R}}}, \tilde{\mathbf{D}} \rangle_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}}}$$

$$s_{\mathrm{b}}^{\mathrm{D05}} = \frac{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{B}}}, \mathbf{D} \rangle_{(\tilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{B}}}, \tilde{\mathbf{D}} \rangle_{(\tilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}} \qquad s_{\mathrm{o}}^{\mathrm{D05}} = \frac{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{R}}}, \mathbf{D} \rangle_{(\tilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}{\langle \hat{\mathbf{u}}_{\tilde{\mathbf{R}}}, \tilde{\mathbf{D}} \rangle_{(\tilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}$$

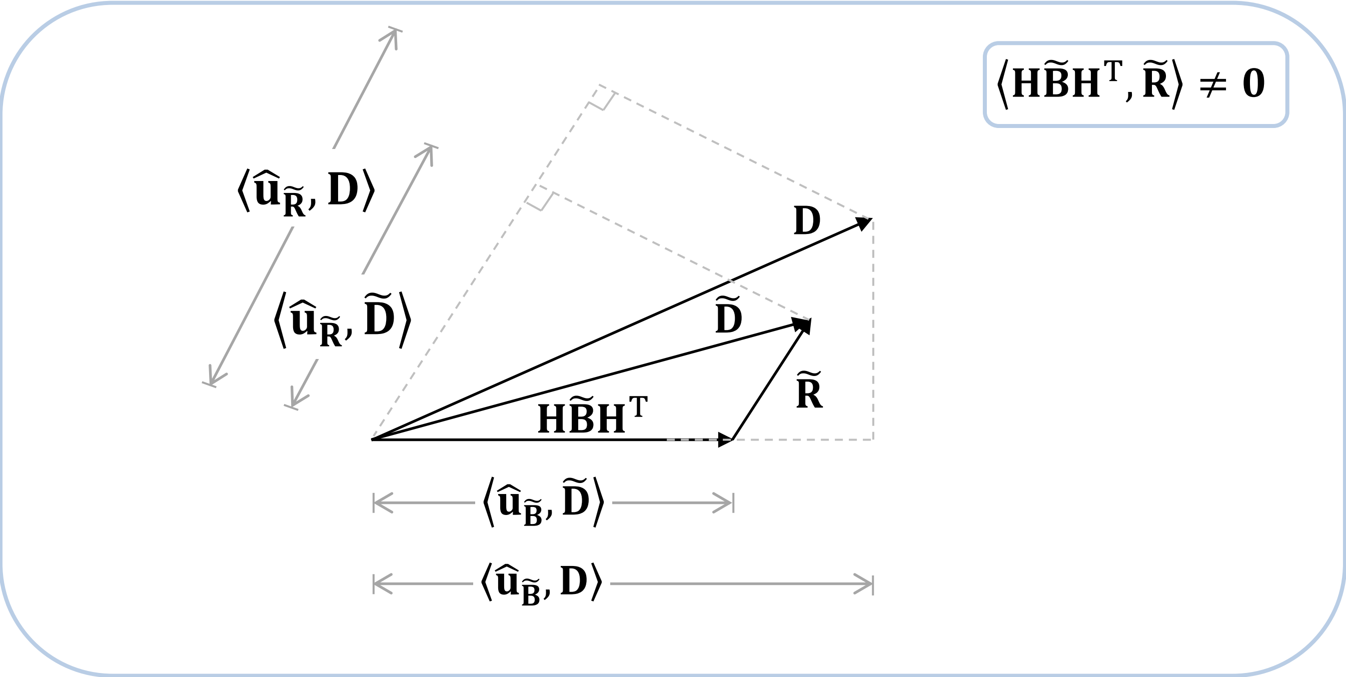


$$\boxed{\langle \mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}, \tilde{\mathbf{R}} \rangle = \mathbf{0}}$$

Geometric Interpretations

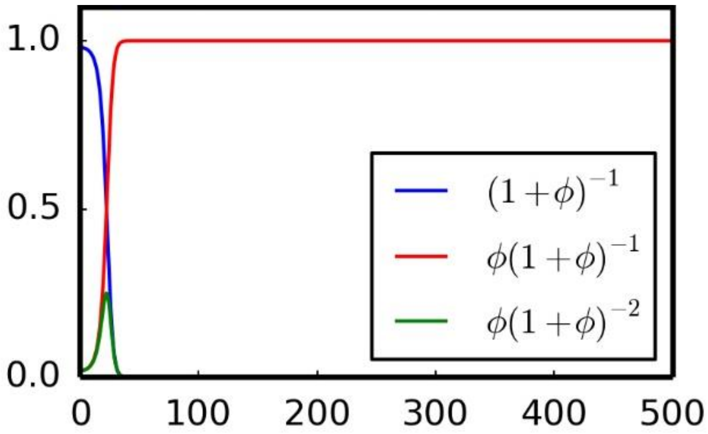
$$s_{\mathrm{b}}^{\mathrm{DI01}} = \frac{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{B}}}, \mathbf{D} \rangle_{(\widetilde{\mathbf{D}} \otimes \widetilde{\mathbf{D}})^{-1}}}{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{B}}}, \widetilde{\mathbf{D}} \rangle_{(\widetilde{\mathbf{D}} \otimes \widetilde{\mathbf{D}})^{-1}}} \qquad s_{\mathrm{o}}^{\mathrm{DI01}} = \frac{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{R}}}, \mathbf{D} \rangle_{(\widetilde{\mathbf{D}} \otimes \widetilde{\mathbf{D}})^{-1}}}{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{R}}}, \widetilde{\mathbf{D}} \rangle_{(\widetilde{\mathbf{D}} \otimes \widetilde{\mathbf{D}})^{-1}}}$$

$$s_{\mathrm{b}}^{\mathrm{D05}} = \frac{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{B}}}, \mathbf{D} \rangle_{(\widetilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{B}}}, \widetilde{\mathbf{D}} \rangle_{(\widetilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}} \qquad s_{\mathrm{o}}^{\mathrm{D05}} = \frac{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{R}}}, \mathbf{D} \rangle_{(\widetilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}{\langle \hat{\mathbf{u}}_{\widetilde{\mathbf{R}}}, \widetilde{\mathbf{D}} \rangle_{(\widetilde{\mathbf{D}} \otimes \mathbf{I})^{-1}}}$$

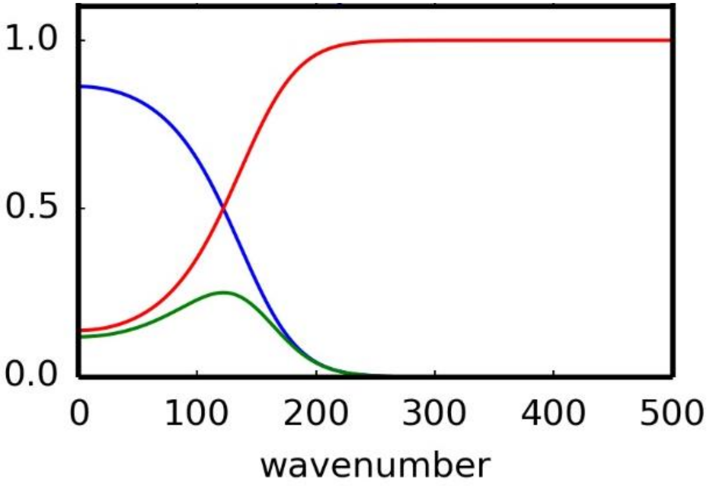


$$\boxed{\langle \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}, \widetilde{\mathbf{R}} \rangle \neq \mathbf{0}}$$

Spectral Distinctiveness of Filters

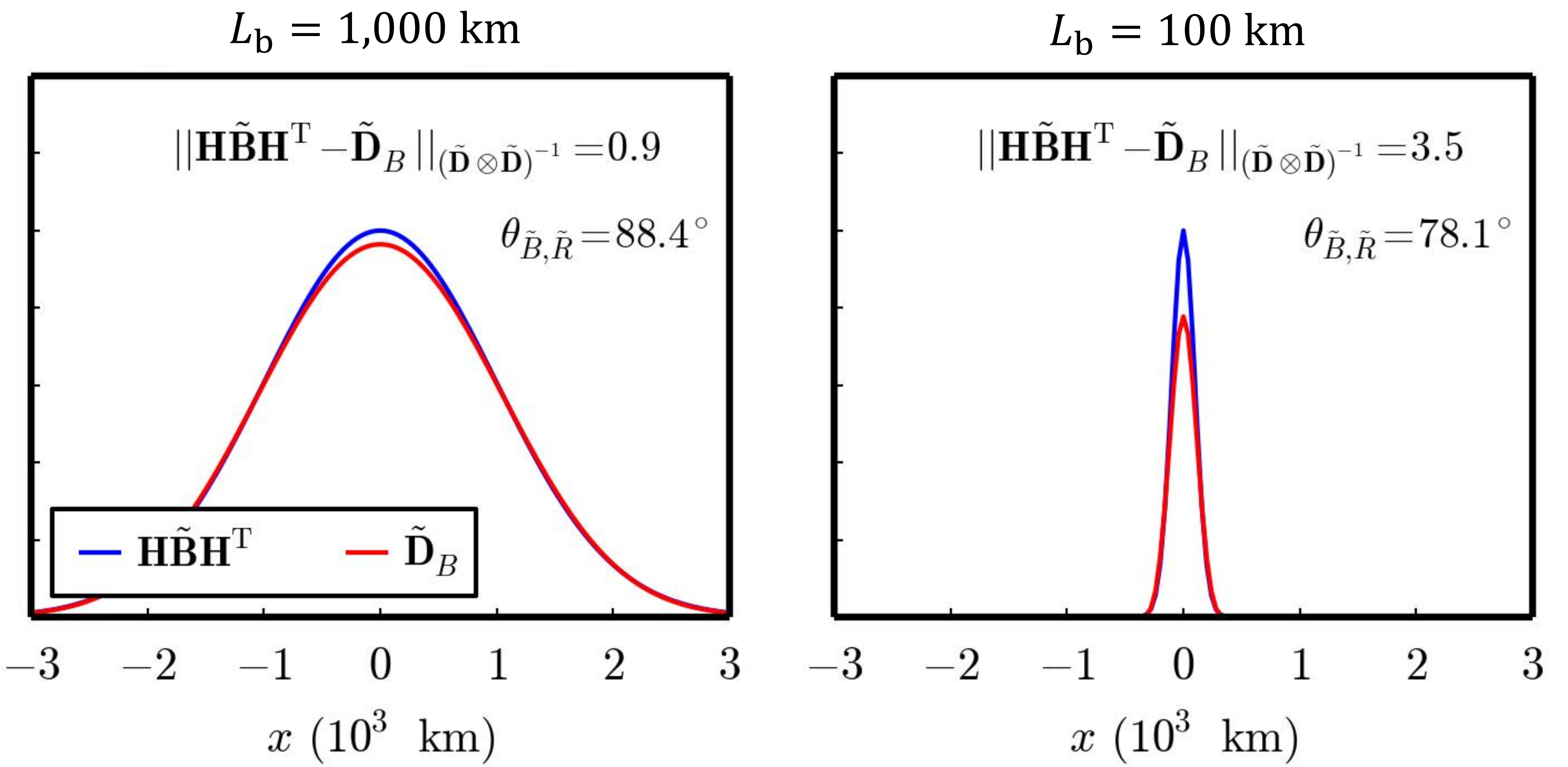

$L_\mathrm{b} = 800$ km

$$\theta_{\tilde{B},\tilde{R}} = 88°$$


$L_\mathrm{b} = 120$ km

$$\theta_{\tilde{B},\tilde{R}} = 78°$$

Error Covariance Filtering Efficiency



$L_\mathrm{b} = 1{,}000$ km

$L_\mathrm{b} = 100$ km

$L = 40{,}000$ km, $\tilde{\sigma}_\mathrm{b}^2 = \tilde{\sigma}_\mathrm{o}^2$, $L_\mathrm{o} = 0$

Statistical Properties

$$\frac{\mathrm{E}[s_{\mathrm{b}}^{\mathrm{DI01}}]}{s_{\mathrm{b}}^{\mathrm{t}}} = 1 + \left(\frac{s_{\mathrm{o}}^{\mathrm{t}}}{s_{\mathrm{b}}^{\mathrm{t}}} - 1\right) \frac{\left\langle \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}, \widetilde{\mathbf{R}} \right\rangle_{(\widetilde{\mathbf{D}}\otimes\widetilde{\mathbf{D}})^{-1}}}{\left\langle \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}, \widetilde{\mathbf{D}} \right\rangle_{(\widetilde{\mathbf{D}}\otimes\widetilde{\mathbf{D}})^{-1}}}$$

$$\frac{\mathrm{E}[s_{\mathrm{o}}^{\mathrm{DI01}}]}{s_{\mathrm{o}}^{\mathrm{t}}} = 1 + \left(\frac{s_{\mathrm{b}}^{\mathrm{t}}}{s_{\mathrm{o}}^{\mathrm{t}}} - 1\right) \frac{\left\langle \mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}, \widetilde{\mathbf{R}} \right\rangle_{(\widetilde{\mathbf{D}}\otimes\widetilde{\mathbf{D}})^{-1}}}{\left\langle \widetilde{\mathbf{R}}, \widetilde{\mathbf{D}} \right\rangle_{(\widetilde{\mathbf{D}}\otimes\widetilde{\mathbf{D}})^{-1}}}$$

$$\mathrm{V}\left[\frac{s_{\mathrm{b}}^{\mathrm{DI01}}}{s_{\mathrm{b}}^{\mathrm{t}}}\right] = \frac{2\mathrm{Tr}\left[\left(\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}\widetilde{\mathbf{D}}^{-1}\left(\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} + \frac{s_{\mathrm{o}}^{\mathrm{t}}}{s_{\mathrm{b}}^{\mathrm{t}}}\widetilde{\mathbf{R}}\right)\widetilde{\mathbf{D}}^{-1}\right)^{2}\right]}{\mathrm{Tr}\left[\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}\widetilde{\mathbf{D}}^{-1}\right]^{2}}$$

$$\mathrm{V}\left[\frac{s_{\mathrm{o}}^{\mathrm{DI01}}}{s_{\mathrm{o}}^{\mathrm{t}}}\right] = \frac{2\mathrm{Tr}\left[\left(\widetilde{\mathbf{R}}\widetilde{\mathbf{D}}^{-1}\left(\frac{s_{\mathrm{b}}^{\mathrm{t}}}{s_{\mathrm{o}}^{\mathrm{t}}}\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}} + \widetilde{\mathbf{R}}\right)\widetilde{\mathbf{D}}^{-1}\right)^{2}\right]}{\mathrm{Tr}\left[\widetilde{\mathbf{R}}\widetilde{\mathbf{D}}^{-1}\right]^{2}}$$

*case where correlations are properly specified

Evaluating Expected Performance

- Case 1:
  - Model error covariances with $L_\mathrm{b} = 1{,}000\ \mathrm{km}$
  - Using observed innovations, calculated $s_\mathrm{b}^\mathrm{DI01} = 2.6$ and $s_\mathrm{o}^\mathrm{DI01} = 1.3$

$\theta_{\tilde{B},\tilde{R}} = 83^\circ$ ⇨

$\mathrm{RMSE}\left[s_\mathrm{b}^\mathrm{DI01}\right] \approx 17\%$

$\mathrm{RMSE}\left[s_\mathrm{o}^\mathrm{DI01}\right] \approx 4\%$

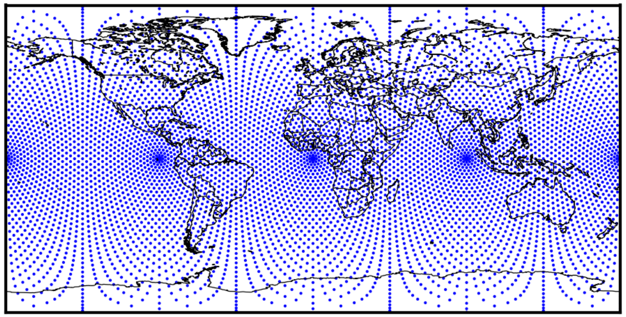
- Case 2:
  - Model error covariances with $L_\mathrm{b} = 250\ \mathrm{km}$
  - Using observed innovations, calculated $s_\mathrm{b}^\mathrm{DI01} = 3.3$ and $s_\mathrm{o}^\mathrm{DI01} = 1.1$

$\theta_{\tilde{B},\tilde{R}} = 26^\circ$ ⇨

$\mathrm{RMSE}\left[s_\mathrm{b}^\mathrm{DI01}\right] \approx 35\%$

$\mathrm{RMSE}\left[s_\mathrm{o}^\mathrm{DI01}\right] \approx 84\%$

Model error covariances $\tilde{\sigma}_\mathrm{b}^2 = \tilde{\sigma}_\mathrm{o}^2$ and $L_\mathrm{o} = 0$



5810-point Lebedev grid, $\Delta x \sim 100\ \mathrm{km} - 330\ \mathrm{km}$

Generalized Algorithm

- DI01 and D05 part of a larger class of algorithm, each defined by their choice of weights

$$\hat{s}_\mathrm{b} = \frac{\mathrm{vec}\left(\widetilde{\mathbf{D}}_B\right)^\mathrm{T}(\mathbf{W}_1 \otimes \mathbf{W}_2)\mathrm{vec}(\mathbf{D}_B)}{\mathrm{vec}\left(\widetilde{\mathbf{D}}_B\right)^\mathrm{T}(\mathbf{W}_1 \otimes \mathbf{W}_2)\mathrm{vec}(\widetilde{\mathbf{D}}_B)} \qquad \hat{s}_\mathrm{o} = \frac{\mathrm{vec}\left(\widetilde{\mathbf{D}}_R\right)^\mathrm{T}(\mathbf{W}_1 \otimes \mathbf{W}_2)\mathrm{vec}(\mathbf{D}_R)}{\mathrm{vec}\left(\widetilde{\mathbf{D}}_R\right)^\mathrm{T}(\mathbf{W}_1 \otimes \mathbf{W}_2)\mathrm{vec}(\widetilde{\mathbf{D}}_R)}$$

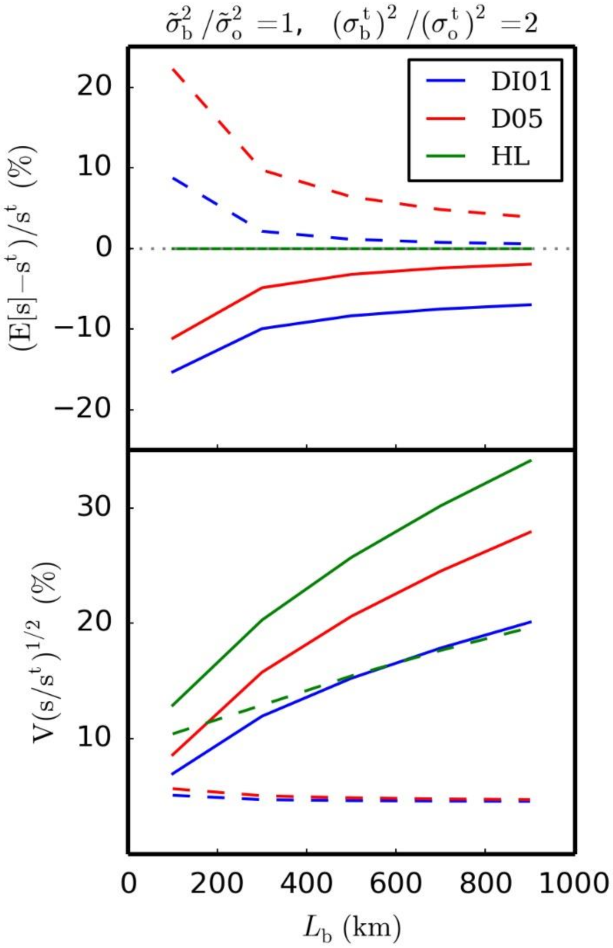
- Weighting of $(\widetilde{\mathbf{D}}_B \otimes \widetilde{\mathbf{D}}_B)^{-1}$ and $(\widetilde{\mathbf{D}}_R \otimes \widetilde{\mathbf{D}}_R)^{-1}$gives algorithm to satisfy the χ^2 diagnostic with $\hat{s}_\mathrm{b} = \hat{s}_\mathrm{o}$

Weighting in Least-Squares Fitting

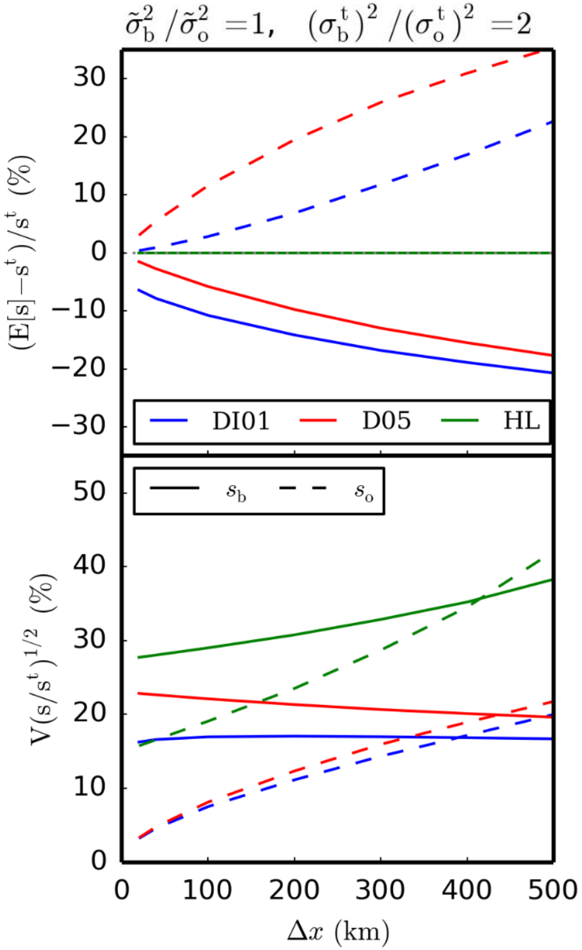
- If $\mathbf{d} \sim \mathcal{N}(\mathbf{0}, \mathbf{D})$, then the sample covariance $\boldsymbol{S}$ follows a Wishart distribution with $V[S_{i,j}] \propto D_{i,j}^2 + D_{i,i}D_{j,j}$

Method	J_B weighting	J_R weighting
DI01	$(\widetilde{\mathbf{D}}_B \otimes (\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}))^{-1}$	$(\widetilde{\mathbf{D}}_R \otimes \widetilde{\mathbf{R}})^{-1}$
D05	$(\widetilde{\mathbf{B}}_B \otimes \mathbf{I})^{-1}$	$(\widetilde{\mathbf{R}}_R \otimes \mathbf{I})^{-1}$
χ^2	$(\widetilde{\mathbf{D}}_B \otimes \widetilde{\mathbf{D}}_B)^{-1}$	$(\widetilde{\mathbf{D}}_R \otimes \widetilde{\mathbf{D}}_R)^{-1}$

Comparisons Between Methods



1D periodic domain
$L = 40{,}000$ km, $\Delta x = 40$ km , $L_{\rm o} = 0$



1D periodic domain
$L = 40{,}000$ km, $L_{\rm b} = 600$ km , $L_{\rm o} = 0$

Conclusions

- DI01, D05, and HL all fit modelled to observation error covariances
  - Fitting is explicit for HL, implicit for DI01 and D05
- Conceptually, DI01 and D05 only differ by the weighting of the cost functions
  - Numerical differences between DI01 and D05 are important
- Performance of DI01 and D05 can be quantified through geometric quantities like $\theta_{\tilde{B},\tilde{R}}$
- Analytic results for error covariances scaling statistics

Sitwell, Michael, and Richard Ménard. "Framework for the comparison of a priori and a posteriori error variance estimation and tuning schemes." *Quarterly Journal of the Royal Meteorological Society* 146.731 (2020): 2547-2575.

Extra Slides

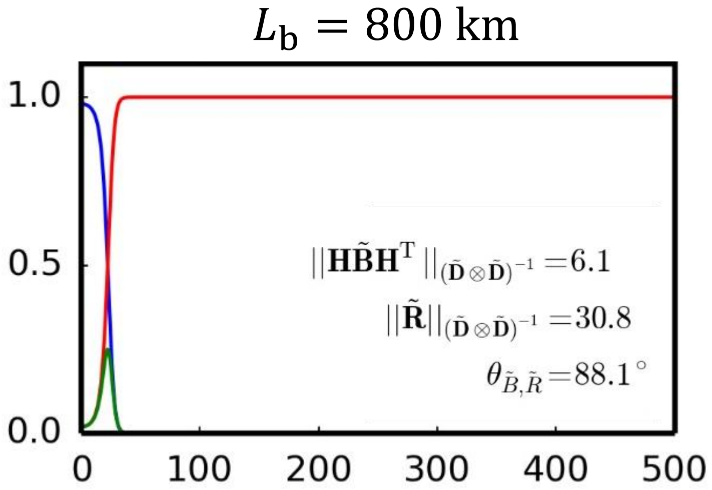
Spectral Distinctiveness of Filters

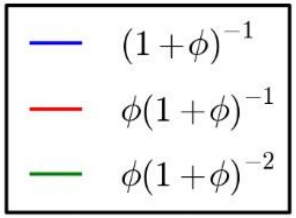
$$\left\|\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}\right\|^{2}_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}} = \mathrm{Tr}[\tilde{\mathbf{F}}\tilde{\mathbf{F}}] = \sum_{i}\frac{1}{(1+\phi_i)^2}$$

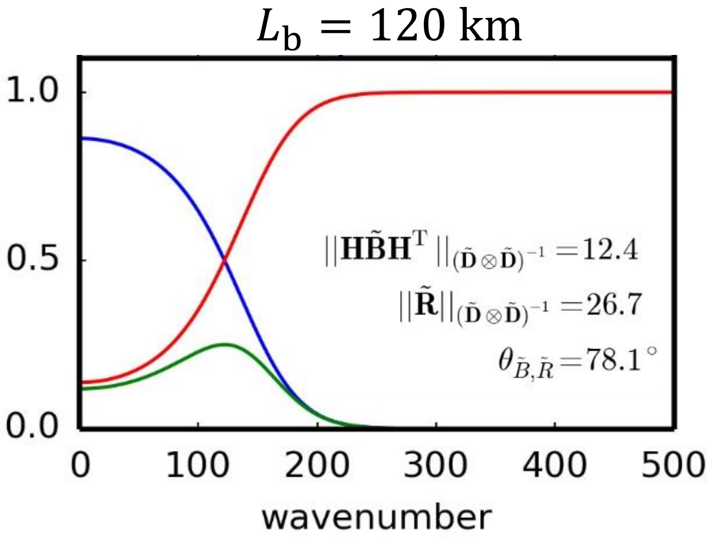
$$\left\|\tilde{\mathbf{R}}\right\|^{2}_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}} = \mathrm{Tr}[(\mathbf{I}-\tilde{\mathbf{F}})(\mathbf{I}-\tilde{\mathbf{F}})] = \sum_{i}\frac{\phi_i^2}{(1+\phi_i)^2}$$

$$\left\langle\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}},\tilde{\mathbf{R}}\right\rangle_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}} = \mathrm{Tr}[\tilde{\mathbf{F}}(\mathbf{I}-\tilde{\mathbf{F}})] = \sum_{i}\frac{\phi_i}{(1+\phi_i)^2}$$

$$\cos\left(\theta_{\tilde{B},\tilde{R}}\right) = \frac{\left\langle\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}},\tilde{\mathbf{R}}\right\rangle_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}}}{\left\|\mathbf{H}\tilde{\mathbf{B}}\mathbf{H}^{\mathrm{T}}\right\|_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}}\times\left\|\tilde{\mathbf{R}}\right\|_{(\tilde{\mathbf{D}}\otimes\tilde{\mathbf{D}})^{-1}}}$$







Limiting Cases

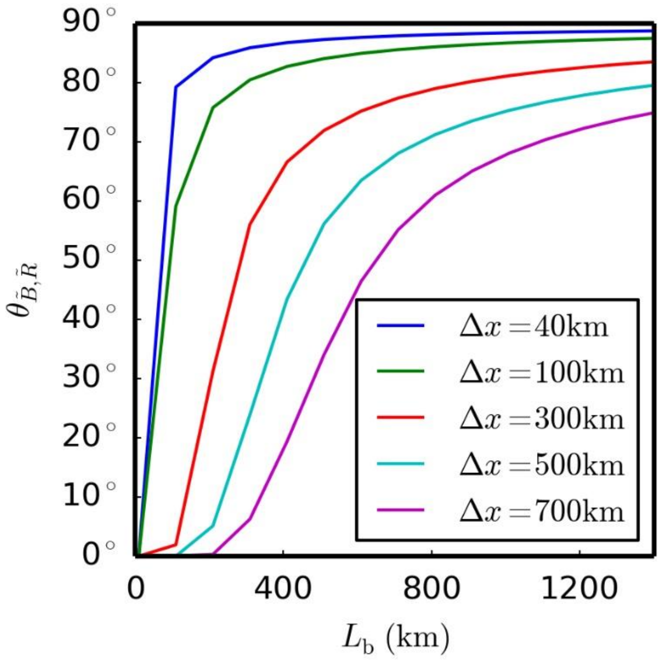
- 1D periodic domain with $L_\mathrm{b} \to \infty$, $L_\mathrm{o} \to 0$:
  - Only one overlapping wavenumber

$$\cos\left(\theta_{\tilde{B},\tilde{R}}\right) = \frac{1}{\sqrt{\left(1+\frac{1}{\phi_0}\right)^2 \left(\frac{L}{\Delta x}-1\right)+1}} \qquad \theta_{\tilde{B},\tilde{R}} \xrightarrow{\Delta x \to 0} 90°$$

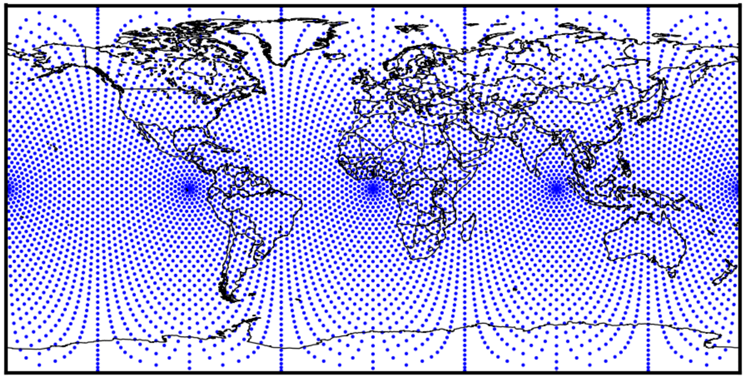
- $\mathbf{H\tilde{B}H}^\mathrm{T} \propto \mathbf{\tilde{R}}$:
  - $\mathbf{\tilde{F}}$ and $\mathbf{I}+\mathbf{\tilde{F}}$ have flat spectra

$$\theta_{\tilde{B},\tilde{R}} = 0°$$

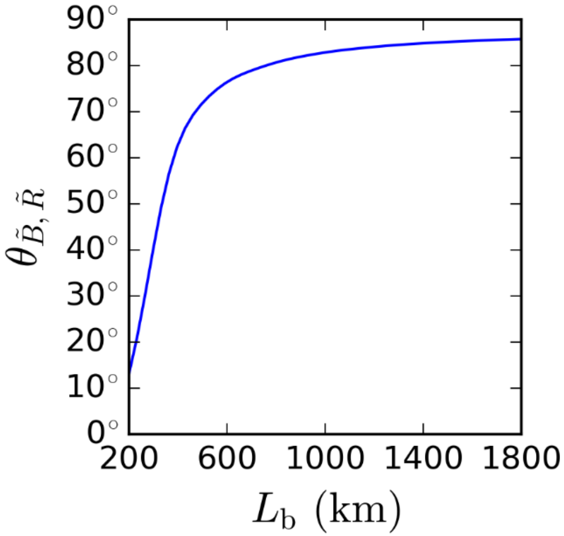
Angles Between Error Covariances



1D periodic domain
$L = 40{,}000$ km, $\tilde{\sigma}_\mathrm{b}^2 = \tilde{\sigma}_\mathrm{o}^2$, $L_\mathrm{o} = 0$

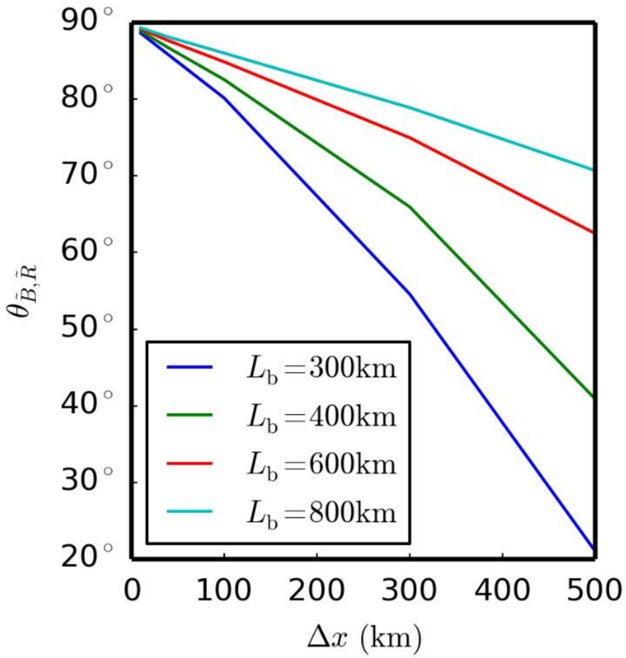


5810-point Lebedev grid, $\Delta x \sim 100$ km $-$ 330 km



$\tilde{\sigma}_\mathrm{b}^2 = \tilde{\sigma}_\mathrm{o}^2$, $L_\mathrm{o} = 0$

Angles Between Error Covariances



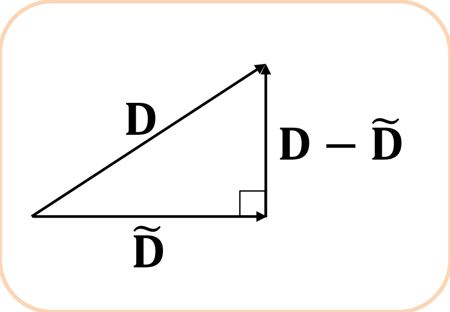
1D periodic domain

$L = 40{,}000$ km, $\tilde{\sigma}_b^2 = \tilde{\sigma}_o^2$, $L_o = 0$

χ^2 Diagnostic

- Weighting of $(\tilde{\mathbf{D}}_B \otimes \tilde{\mathbf{D}}_B)^{-1}$ and $(\tilde{\mathbf{D}}_R \otimes \tilde{\mathbf{D}}_R)^{-1}$gives algorithm to satisfy the χ^2 diagnostic with $\hat{s}_\mathrm{b} = \hat{s}_\mathrm{o}$

$$\chi^2 \text{ diagnostic satisfied if } \left\langle \tilde{\mathbf{D}}, \mathbf{D} \right\rangle_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}} = \left\| \tilde{\mathbf{D}} \right\|^2_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}} = N_\mathrm{obs}$$



$$\left\| \mathbf{D} - \tilde{\mathbf{D}} \right\|^2_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}} = \left\| \mathbf{D} \right\|^2_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}} - \left\| \tilde{\mathbf{D}} \right\|^2_{(\tilde{\mathbf{D}} \otimes \tilde{\mathbf{D}})^{-1}}$$

Comparing to Hollingsworth and Lönnberg

- First minimize $J_B^{\mathrm{HL}} = \frac{1}{2}\left\| s_\mathrm{b}\mathbf{V}\circ\widetilde{\mathbf{D}} - \mathbf{V}\circ\mathbf{D}\right\|^2_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2)^{-1}}$

minimize w.r.t. S_b

$$s_\mathrm{b}^{\mathrm{HL}} = \frac{\left\langle \mathbf{V}\circ\widetilde{\mathbf{D}}, \mathbf{V}\circ\mathbf{D}\right\rangle_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2)^{-1}}}{\left\|\mathbf{V}\circ\widetilde{\mathbf{D}}\right\|^2_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{b}^2)^{-1}}}$$

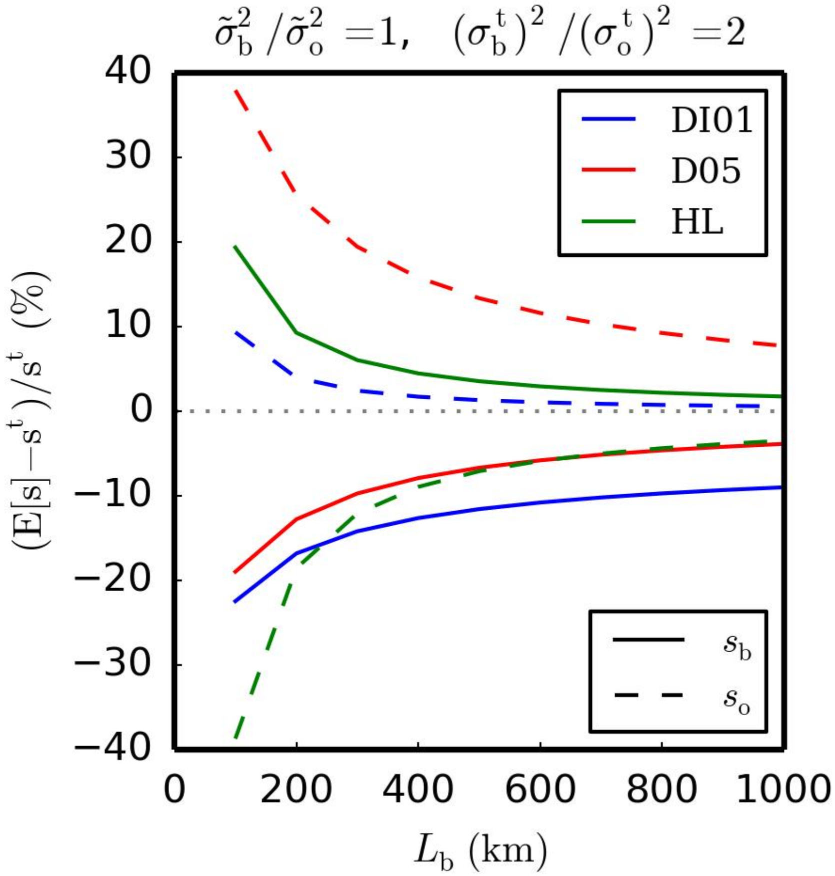
$$\mathbf{V} \equiv \begin{bmatrix} 0 & 1 & 1 & \\ 1 & 0 & 1 & \cdots \\ 1 & 1 & 0 & \\ & \vdots & & \ddots \end{bmatrix}$$

- Then minimize $J_R^{\mathrm{HL}} = \frac{1}{2}\left\| \mathbf{I}\circ\left(s_\mathrm{b}^{\mathrm{HL}}\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T} + s_\mathrm{o}\widetilde{\mathbf{R}}\right) - \mathbf{I}\circ\mathbf{D}\right\|^2_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2)^{-1}}$

minimize w.r.t. S_o

$$s_\mathrm{o}^{\mathrm{HL}} = \frac{\left\langle \mathbf{I}\circ\widetilde{\mathbf{R}}, \mathbf{I}\circ\left(\mathbf{D} - s_\mathrm{b}^{\mathrm{HL}}\mathbf{H}\widetilde{\mathbf{B}}\mathbf{H}^\mathrm{T}\right)\right\rangle_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2)^{-1}}}{\left\|\mathbf{I}\circ\widetilde{\mathbf{R}}\right\|^2_{(\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2\otimes\widetilde{\boldsymbol{\Sigma}}_\mathrm{o}^2)^{-1}}}$$

Comparisons Between Methods



1D periodic domain
$L = 40{,}000$ km, $\Delta x = 40$ km , $L_o = 40$ km